ELSEVIER

Comparability of National Death Index *Plus* and Standard Procedures for Determining Causes of Death in Epidemiologic Studies

MICHELE MORIN DOODY, MS, HOWARD M. HAYES, DVM, AND ROBERT BILGRAD, MA, MPH

PURPOSE: To determine whether causes of death obtained through National Death Index (NDI) *Plus* are comparable to those obtained by requesting death certificates from state vital statistics offices and having deaths coded by contractor nosologists.
METHODS: The authors compared underlying cause of death codes obtained from NDI *Plus* with those assigned by contractor nosologists for a sample of 250 known decedents.
RESULTS: The underlying cause of death codes differed for 18 (7%) of 249 successful matches. Independent coding by an expert National Center for Health Statistics (NCHS) nosologist trainer revealed that seven of these had an NDI *Plus* code that matched the code provided by the NCHS nosologist and a contractor nosologist code that did not match the NCHS nosologist code, seven had a contractor nosologist code that matched the NCHS nosologist code and an NDI *Plus* code that did not match the NCHS nosologist code, and four had both an NDI *Plus* and a contractor nosologist code that did not match the NCHS nosologist code. The level of disagreement with the NCHS nosologist and the organ systems involved were similar for NDI *Plus* and the contractor nosologist.
CONCLUSIONS: The authors report that NDI *Plus* provides comparable information within a substantially shorter time period for most states and, for known decedents, at about half the cost of standard procedures.
Ann Epidemiol 2001;11:46–50.

KEY WORDS: Cause Of Death, Death Certificates, Epidemiologic Methods.

INTRODUCTION

The National Death Index (NDI), a service of the United States National Center for Health Statistics (NCHS), assists health and medical researchers in determining the vital status of study subjects (1). The NDI is a national database containing identifying death record information submitted annually since 1979 by all state vital statistics offices, the District of Columbia, New York City, Puerto Rico, and the Virgin Islands. Searches against the NDI file provide the location and date of death, and the death certificate number for each decedent. In March 1997, the NCHS announced a new service, NDI *Plus* (2), which, for a modest additional fee, provides the underlying, contributing, and all other causes of death coded to the International Classification of Diseases, 9th Revision (ICD-9) (3). Causes of death are obtained via linkage with the NCHS Mortality Statistics file which is also compiled from records submitted by the state vital statistics offices. While this approach is substantially less expensive and time-consuming than standard epidemiologic procedures for obtaining causes of death for known decedents, comparability of codes obtained from NDI *Plus* and contractor nosologists is of concern for long-term follow-up studies. In this paper, underlying cause of death codes derived from the two sources are directly compared for a random sample of known decedents within a nationally-distributed cohort under study by the National Cancer Institute.

MATERIALS AND METHODS

The study cohort is comprised of approximately 145,000 individuals who were certified by the American Registry of Radiologic Technologists (ARRT) for at least two years between 1926 and 1982 (4). Linkage with mortality records from the Social Security Administration (SSA), NDI, state vital statistics offices, and other public records identified a total of 6831 deaths between 1979 and 1994. Standard procedures for determining causes of death include requesting death certificates from state vital statistics offices

From the Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD (M.M.D., H.M.H.); and the National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, MD (R.B.).
Address reprint requests to: Michele M. Doody, Radiation Epidemiology Branch, National Cancer Institute, Executive Plaza South, Room 7088, Bethesda, MD 20892–7362.
Authors are listed in the order of their contribution.
Received January 12, 2000; accepted June 28, 2000.

655 Avenue of the Americas, New York, NY 10010
1047-2797/01/$–see front matter
PII S1047-2797(00)00177-0

Selected Abbreviations and Acronyms

NDI = National Death Index
NCHS = National Center for Health Statistics
ICD-9 = International Classification of Diseases, 9th Revision
ARRT = American Registry of Radiologic Technologists
SSA = Social Security Administration

and having the underlying causes of death coded by trained contractor nosologists. The routine coding protocol involves independent coding by two nosologists according to ICD-9 (3), with arbitration of any conflicting codes by a medical coding supervisor.

A sample of 250 decedents was selected from among all decedents for submission to NDI *Plus*. The sample was selected randomly within gender group (175 female and 75 male decedents), with the size of the gender groups determined by the distribution of the overall cohort (73% female). The underlying causes of death provided by NDI *Plus* were compared to those previously coded by the contractor nosologists. Level of agreement was defined by priority as: match on all four digits; match on first three digits only; match within same major organ system only; or no match (different major organ systems). Discrepancies for most cancers and for other causes of death will generally be minor for a match on three digits only, however, such a match will occasionally result in a major discrepancy (e.g., leukemia subtypes). Major organ system categories (see Table 1) were chosen to be consistent with mortality rate files developed by Dr. Richard Monson (5), and updated and revised by Information Management Services, Inc. for use in the Observed *versus* Expected Events (O/E) software program (6). These rate files allow for analyses of deaths occurring as early as 1925. Death certificates for which the underlying cause of death codes disagreed at any level were blindly coded by an expert NCHS nosologist trainer (Julia Raynor).

RESULTS

Two hundred and forty-nine of the 250 technologists submitted linked successfully with NDI *Plus*. The one subject who did not link was initially found through an NDI match in 1990 and was confirmed deceased by death certificate. Subsequent tracing yielded a slightly different, unfortunately erroneous Social Security number (SSN) for this subject. When submitted the second time, the combination of an inexact SSN and a common surname precluded an eligible NDI *Plus* match. Due to the sampling scheme, there were proportionately more females in the sample than in the group of all decedents (70% *versus* 59%). Distributions in the sample by race, year of birth, year of death, age at death, and state of death closely reflected those among all decedents. The sample was predominantly white (89%), and 58% were born before 1930 (range, 1892–1959). Age at death spanned between 23 and 97 years (mean, 61), with a mean year of death of 1987. Deaths occurred in 44 states and the District of Columbia.

A total of 231 (93%) of the underlying cause of death codes obtained from NDI *Plus* and the contractor nosologists matched exactly on all four digits (Table 1). Most deaths were from malignant neoplasms (35%) and diseases of the circulatory system (32%). Two AIDS-related deaths did not match initially. The NDI *Plus* codes reflected the rules in effect when the deaths occurred (1985 and 1986), whereas the contractor and NCHS expert codes were assigned after the introduction of the AIDS classification in 1987. When the contractor and NCHS nosologists re-coded using the same rules, codes from all three sources matched exactly.

Of the 18 codes that differed between NDI *Plus* and the contractor, three (17%) matched on the first three digits (two external causes and one respiratory system), eight others (44%) were within the same major organ system (three circulatory, two cancer, two external, and one respiratory), and seven (39%) differed according to major organ system (Table 2). Arbitration of discrepant codes by an expert NCHS nosologist trainer revealed that seven of these had an NDI *Plus* code that matched the code provided by the NCHS nosologist and a contractor nosologist code that did not match the NCHS nosologist code; seven had a contractor nosologist code that matched the NCHS nosologist code and an NDI *Plus* code that did not match the NCHS nosologist code; and four had both an NDI *Plus* and a contractor nosologist code that did not match the NCHS nosologist code. The discrepancies were similar according to their level of agreement with the NCHS nosologist and the major organ systems involved (Table 3).

A comparative evaluation was made of the cost and level of effort required for the two methods. Tasks required by both included staff management and supervision (NDI *Plus*, 20 hrs *versus* contractor, 26 hrs), computer programming (18 hrs *versus* 8 hrs), and clerical review (20 hrs *versus* 88 hrs). Additionally, the standard method required secretarial support for contacting state vital statistics offices (8 hrs), nosology coding of causes of death (67 hrs), and double-key data entry for cause of death codes (4 hrs). The fee for the NDI *Plus* submission was less than half that for acquiring death certificates ($1,278 *versus* $2,750). All factors considered, the estimated per subject cost to obtain underlying causes of death through NDI *Plus* was about half that of our standard approach ($20.34 *versus* $39.26).

Initial approvals from NDI and the individual states were obtained for this study many years ago, thus, we cannot estimate the length of time it would take to obtain results for new studies. For previously approved studies, it takes about two weeks to obtain repeat approval and a returned

TABLE 1. Distribution of underlying causes of death for codes that matched on all four digits between the contractor nosologists and NDI *Plus*

Major organ system (ICD-9)[a]	Number of decedents	%
Infectious and parasitic diseases (001–139)	11	5
Malignant neoplasms (140–208, 238.4, 238.6, 289.8, except 202.2, 202.3, 202.5, 202.6)	81	35
Digestive	13	
Lung	21	
Breast	18	
Genitourinary	9	
Endocrine, nutritional and metabolic diseases, and immunity disorders (240–279)	6	3
Diseases of blood and blood-forming organs (280–289, except 289.8)	2	1
Mental disorders (290–319)	2	1
Diseases of the nervous system and sense organs (320–389)	3	1
Diseases of the circulatory system (390–459)	74	32
Ischemic heart disease	45	
Other forms of heart disease	15	
Cerebrovascular disease	11	
Diseases of the respiratory system (460–519)	17	7
Diseases of the digestive system (520–579)	11	5
Diseases of the genitourinary system (580–629)	2	1
Diseases of the musculoskeletal system and connective tissue (710–739)	2	1
Congenital anomalies (740–759)	1	<1
Symptoms, signs, and ill-defined conditions (780–799)	1	<1
Injury and poisoning (800–999)	17	7
Other[b]	1	<1
TOTAL Matches[c]	231	100

[a] International Classification of Diseases, 9th Revision (WHO 1977).
[b] Malignant histiocytosis (202.3).
[c] There were no deaths from benign neoplasms (210–239, except 237.7, 238.4–238.7), complications of pregnancy, childbirth, and the puerperium (630–676), diseases of the skin and subcutaneous tissue (680–709), or perinatal conditions (760–779).

file from NDI. The length of time it takes to obtain death certificates from state health departments is variable, and can depend upon changes in state regulations, staff changes within state vital statistics offices, and the established relationship between the requestor and the states. Based on a state-by-state review of their most recent submissions for several of our ongoing studies and covering all states except Hawaii, our contractor reports receiving certificates from 18 states in less than one month, 16 states in about one month, six states in about two months, and nine states in 3–6 months (six months for New York State, 11 months for New York City).

DISCUSSION

We are aware of only one prior study in which causes of death obtained through NDI *Plus* were evaluated. Among 493 decedents in a cohort of New Jersey chemical plant workers, Sathiakumar and colleagues (7) reported discrepancy rates of 4% for NDI *Plus* codes, 4% for the final study codes, and 6–7% for the study nosologist's original codes when compared to the reference standard provided by an NCHS nosologist. Since our primary goal was to evaluate the comparability of codes between the two sources, we did not pursue NCHS nosologist coding of deaths that matched on all four digits, thus, we cannot estimate discrepancy rates for our sample.

For the vast majority of deaths (93%) in this sample, the cause of death codes obtained through the contractor nosologists and NDI *Plus* matched exactly. Among the 18 codes that were dissimilar, seven codes from each source matched the NCHS nosologist code exactly and 11 codes from each source disagreed at some level. There were no differences in the level of disagreement or the major organ systems involved. We conclude that the cause of death information provided by NDI *Plus* is comparable to that determined through our standard procedures. The nationwide distribution of deaths among our sample provides added confidence in this finding.

A major disadvantage of using NDI *Plus* for epidemiologic studies is the unavailability of computerized records for deaths occurring before 1979. Another drawback is that follow-back information on hospitals and physicians is not available for pursuing medical record validation of causes of death (7).

Advantages to NDI *Plus* include comparability of causes of death with national mortality rate files and availability of coded contributing causes of death for use in internal

TABLE 2. Discrepancies between the contractor nosologists and NDI *Plus* underlying cause of death codes, and results of blinded coding by an expert National Center for Health Statistics (NCHS) nosologist trainer

Subject number	Year of death	State of death	Causes of death from death certificate (Part I)	ICD9[a] Codes: Contractor nosologists	ICD9[a] Codes: NDI *Plus*	ICD9[a] Codes: NCHS nosologist
3rd digit agreement						
1	1991	AL	Thermal burns Two vehicle collision	E8121	E8129	E8129
2	1987	AR	Trauma to head and body Auto wreck	E8129	E8120	E8129
3	1993	CA	Acute cardiorespiratory arrest Hepatic encephalopathy Cirrhosis	5715	5712	5715
Same major organ system[b]						
4	1981	IL	Cardiopulmonary arrest Congestive heart failure Arteriosclerotic cardiovascular disease	4148	412	412
5	1988	IN	Acute pulmonary edema Coronary insufficiency Athersclerosis	411	4148	4148
6	1984	NC	Metastatic adenocarcinoma	1940	1991	1991
7	1993	TX	Squamous cell carcinoma face Brain malignancy, unspecified Basal cell carcinoma	1733	1919	1733
8	1982	OR	Pulmonary embolus Pancreatitis Common duct stone	5770	5745	5770
9	1984	GA	Carbon monoxide intoxication	E867	E8683	E8683
10	1980	NYC	Cardiopulmonary arrest End stage mitral valve disease Coronary artery disease	4240	4149	3949
11	1987	NYC	Smoke inhalation and thermal burns Fire at (home address) Pending Fire Marshal's investigation	E9881	E8912	E8902
Different major organ system[b]						
12	1993	IA	Acute myocardial infarction End stage lung cancer	1629	410	1629
13	1980	SD	Bacterial pneumonia Chronic aspiration Chronic coma	4829	436	438
14	1982	OR	Aspiration of vomitus Gastroenteritis	558	4871	558
15	1986	TX	Acute myocardial infarction Left renal kidney transplant rejection	5939	2503	2503
16	1982	WA	Cardiac arrest Aspiration pneumonia into lung Pulmonary edema	7999	5327	5327
17	1987	TX	Aspiration of gastric contents and blood Fatty metamorphosis of liver, severe Multiple laceration of face and head, fractured nose	E888	5718	E9289
18	1984	WV	Adult respiratory distress syndrome Sepsis due to candida albicans Surgery for resection of stomach	5379	1125	5379

[a] International Classification of Diseases, 9th Revision (WHO 1977).
[b] See Table 1 for definition of major organ systems.

TABLE 3. Highest level of agreement with an expert NCHS nosologist trainer for underlying causes of death that differed between the contractor nosologists and NDI *Plus*

Highest level of agreement with NCHS nosologist	Contractor nosologists		NDI *Plus*	
	Number	%	Number	%
4th digit	7	39	7	39
3rd digit	1	6	2	11
Same major organ system[a]	6	33	5	28
Malignant neoplasms	1		1	
Circulatory system	3		2	
Other	4		2	
Different major organ system[a,b]	4	22	4	22
Malignant neoplasms	0		1	
Circulatory system	1		0	
Other	3		3	
TOTAL Discrepancies	18	100	18	100

[a] See Table 1 for definition of major organ systems.
[b] Major organ system as determined by NCHS nosologist.

cohort comparisons (7). For deaths occurring in 1979 or later, NDI *Plus* provides comparable cause of death data within a substantially shorter time period for most states. NDI *Plus* is especially advantageous for multi-center and nationwide studies in which the task of obtaining death certificates from many different states can be quite cumbersome. Finally, for known decedents, the cost of obtaining cause of death codes through NDI *Plus* is about half that of standard epidemiologic procedures.

This research was supported by contract N01-CP-33013 with the National Cancer Institute. The authors gratefully acknowledge Julia Raynor of the National Center for Health Statistics for expert nosology coding, Kathleen Chimes and Carole Prusinowski of Westat, Inc. for data coordination, Peggy Bell of Westat, Inc. for nosology supervision, Bob Banks of Information Management Services, Inc. for biomedical computing, and Ken Kochanek of NCHS for helpful comments on the manuscript.

REFERENCES

1. Bilgrad R. National Death Index User's Manual. Hyattsville, MD: National Center for Health Statistics, Centers for Disease Control and Prevention; 1995.
2. Bilgrad R. National Death Index *Plus*: Coded Causes Of Death Supplement to the National Death Index User's Manual. Hyattsville, MD: National Center for Health Statistics, Centers for Disease Control and Prevention; 1997.
3. World Health Organization. Manual of the International Classification of Diseases, Injuries, and Causes of Death, 9th Revision. Geneva, Switzerland; 1977.
4. Doody MM, Mandel JS, Lubin JH, Boice JD Jr. Mortality among U.S. radiologic technologists, 1926–1990. Cancer Causes Control. 1998; 9:67–75.
5. Monson RR. Analysis of relative survival and proportional mortality. Comput Biomed Res. 1974; 7:325–332.
6. Helde TT Jr. O/E System: Observed Versus Expected Events. User's Guide. Version 3.8. Rockville, MD: Information Management Services, Incorporated; 1991.
7. Sathiakumar N, Delzell E, Abdalla O. Using the National Death Index to obtain underlying cause of death codes. J Occup Environ Med. 1998; 40:808–813.